

TARTU ÜLIKOOL
LOODUS- JA TÄPPISTEADUSTE VALDKOND
MATEMAATIKA JA STATISTIKA INSTITUUT

Kalder Maarand

Statistilise kompetentsi vajadus Eestis tegutsevates ettevõtetes

Bakalaureusetöö (9 EAP)

Juhendajad: lektor Mare Vähi

Silva Kasela, PhD

Kristi Läll, MSc

Tartu
2017

Statistilise kompetentsi vajadus Eestis tegutsevates ettevõtetes

Bakalaureusetöö

Kalder Maarand

Lühikokkuvõte: Käesoleva bakalaureusetöö eesmärgiks on saada ettekujutus statistikute ja analüütikute vajadusest ja rakendamisest Eesti ettevõtetes. Töö raames viidi läbi küsitlus, mis paluti täita ettevõtte tööülesannete või personaliostuste eest vastutaval isikul. Vastuseid koguti aprillis ja mais 2017. aastal. Töö esimeses peatükis tutvustatakse analüütikal põhinevate äriliste otsuste ajalugu ja rolli tänapäeva maailmas koos erinevate näidetega, teises peatükis kirjeldatakse valimi moodustamist, küsimustikku ning uuringu läbiviimist, kolmas peatükk tegeleb uuringu tulemuste analüüsiga.

CERCS teaduseriala: P160 Statistika, operatsioonianalüüs, programmeerimine, finants- ja kindlustusmatemaatika

Märksõnad: turu-uuring, statistiline kompetents

The demand for statistical competence in Estonian businesses

Bachelor's thesis

Kalder Maarand

Abstract: The purpose of this bachelor's thesis is to get an understanding of the demand for and applications of statisticians and analysts in Estonian businesses. For this, a survey was carried out to be answered by the person in charge of personnel decisions or work assignment. Responses were collected from April through May 2017. The first chapter briefly introduces some of the history of analytics-based business decisions and describes the current role of analytics in business, with a few examples. The second chapter describes the sample selection, the questionnaire and the process of carrying out the survey, the third chapter contains the analysis of the survey results.

CERCS research specialisation: P160 Statistics, operations research, programming, actuarial mathematics

Keywords: survey, statistical competence

Sisukord

Statistilise kompetentsi vajadusest.....	4
Valim, küsimustik ja uuringu läbiviimine	8
Kirjeldav analüüs	10
Statistilise kompetentsi kohta käivad küsimused.....	10
Vastanute taustinfo.....	15
Koolituse kohta kogutud info.....	18
Kokkuvõte.....	19
Kasutatud kirjandus	20
Lisad.....	22
Lisa 1. Vabavormis vastused	22
Lisa 2. Küsimustik	26

Statistilise kompetentsi vajadusest

Ettevõtete eesmärgiks on teenida võimalikult palju kasumit. Statistiline analüüs saab aidata kasumlikkust maksimiseerida, näiteks läbi tootlikuse tõstmise, kulude vähendamise või tarbija vajaduste parema mõistmise. McKinsey & Company poolt läbiviidud uuring [1] on näidanud, et ettevõtted, kes kasutavad andmeanalüüsi äriliste otsuste vastuvõtmisel, on 5-6% kõrgema tootlikkuse ja kasumlikkusega ning andmeanalüüsil põhinevad turundusotsused tõstavad turundusinvesteeringute tootlikkust 15-20% . Ilmselt mõistavad ka ettevõtted andmeanalüüsi kasulikkust, sest nõudlus analüütikute järele on viimase paarikümne aastaga tõusnud märgatavalt: American Statistical Association veebilehel avaldatud artikli [2] põhjal töötas 2014. aastal Ameerika Ühendriikides statistiku ametikohal 85 000 inimest, kusjuures 2010. aastal oli see arv 28 000. Kuni 2022. aastani ulatuva prognoosi kohaselt kasvab statistika-alaste ametikohtade arv 27%, mis on ligikaudu 2,5 korda rohkem kui üle kõikide töövaldkondade prognoositav 11% ametikohtade kasv. See arv võib veel konservatiivne olla, sest see ei sisalda mitmeid töökohti, mis vajavad laialdasi statistikaalaseid teadmisi. Kasvanud on ka ülikoolide arv, mis pakuvad statistikakraade: bakalaureusekraade pakkuvate ülikoolide arv on aastatel 2003-2014 suurenenud 50% ja magistrikraade pakkuvate ülikoolide arv 20%. Kõrge nõudlus peegeldub ka valdkonna palganumbrites: 2014. aastal oli statistiku keskmine palk 84 010 dollarit, võrdluseks oli Ühendriikide keskmine palk 2015. aastal 44 510 dollarit [3]. Ärilise suunitlusega sotsiaalvõrgustiku LinkedIni avaldatud andmete kohaselt [4] oli statistiline analüüs ja andmekaeve pilve- ja hajutatud andmetöötuse (*cloud and distributed computing*) järel teine kõige nõutum oskus töökandidaatidele 2016. aastal ning andmete visualiseerimine kaheksas.

Society For Human Resource Managementi poolt 2016. aastal läbi viidud uuringu [5] kohaselt on neljal ettevõttel viiest andmeanalüüsiga tegelev töötaja ja 72% asutustest oli palganud andmeanalüütiku viimase aasta jooksul. Lisaks, 65% ettevõtetest on suurendanud andmeanalüüsiga tegelevate töökohtade arvu viimase viie aasta jooksul ning 59% kavatsevad suurendada seda arvu järgmise viie aasta jooksul. Suurt nõudlust kinnitab ka asjaolu, et 78% ettevõtetest hindasid sobiliku kandidaadi leidmist andmeanalüütiku kohale keeruliseks. Nõutavaimad oskused kandidaatidele olid andmete tõlgendamine ja esitamine, tabelipõhine andmeanalüüs (nt. Excel), statistilise tarkvara kasutamisoskus (R, SAS, SPSS jne.), suurandmetega seotud programmeerimisoskus ning andmekaeve ja masinõpe.

Hiljutise suurenenud nõudluse (võrreldes umbes 20 aasta taguse ajaga) põhjuseks on tehnoloogia areng ja interneti levik. Andmeid on kogutud ja analüüsitud ärilistel eesmärkidel juba pikka aega, näiteks juba 17. sajandil loodi Inglismaal esimene elutabel ehk suremustabel, millel põhineb ka tänapäevane elukindlustusmatemaatika. Andmed olid lihtsad: surnud inimeste arv ja nende vanus, kuid sellest piisas, et oleks võimalik elukindlustusmaks suurust hinnata palju täpsemalt, kui seda oli seni tehtud. Analüüsipõhised äriotsused hakkasid rohkem levima 20. sajandi teises pooles arvutite levikuga seoses, tänapäeval on andmeid kordades rohkem ning 1996. aastaks olid digitaalsed andmekandjad arenenud kõige odavamaks andmetalletamise meetodiks [6]. Digitaalne hoiustus võimaldab talletada palju rohkem andmeid ning tagab ka kiirema ja mugavama juurdepääsu. Nutitelefonid ja muud värvivõrgu seaded tekitavad järjest enam andmeid: 2012. aastal hinnati, et iga päev luuakse 2,5 eksabaiti andmeid (1 eksabait ehk 1 000 000 terabaiti). Loomulikult on vaja, et keegi neid andmeid korrastaks ja analüüsiks.

Selle suure ja kiirelt kasvava andmehulga kirjeldamiseks on tänapäeval kasutusel väljend *big data* ehk suurandmed, mis esines esmakordselt 1990. aastate lõpus. Selle levikuga on omakorda seotud termin *data scientist* ehk andmeteadlane, kes töötab eelkõige suurandmetega. Leidub erinevaid arvamusi selle kohta, kas statistiku ja andmeteadlase ametid on üldse erinevad, Nate Silver on öelnud, et andmeteadlane on lihtsalt uhkema kõlaga väljend statistiku jaoks [7]. Gil Press [8] kirjutas 2013. aastal, et andmeteadus pole selgelt piiritletud ning uued andmeteaduse programmid kõrgkoolides pakuvad sisuliselt samu programme eelnevalt *business intelligence* nime all. Kui aga nüüd paari viimase aasta jooksul ongi andmeteadlust piiritletud, siis see enamasti hõlmab teadmisi statistikast, matemaatikast, arvutiteadusest ja masinõppest. Teisisõnu, vaja on oskusi, mis võimaldavad andmed kätte saada, töödelda ja korrastada, neid seejärel analüüsida ja viimaks avastusi ja järeldusi esitleda. Tundub, et andmeteadlast peetakse statisiku alaliigiks ning neid eraldab n-ö traditsioonilisest statistikust oskus kirjutada algoritme, mis suudavad reaal-ajas saadavaid andmeid analüüsida ja kiirelt vajaliku info eraldada. Lisaks peaks andmeteadlasel olema põhjalikud teadmised ettevõtte tegevusalast ja tootest. 2016. aastal läbi viidud uuringu [9] kohaselt oli 28% protsendil andmeteadlastel akadeemiline kraad matemaatika ja statistika vallast, inserikraadiga oli 18% ja arvutiteaduse kraadiga 17% andmeteadlastest. Kõige enam leidsid andmeteadlased rakendust tehnoloogiasektoris, kus töötas 46% andmeteadlastest, teised rohkem levinud harud olid turundusteenused (12%) ja finantsteenused (10%).

Tehnoloogiasektori suur populaarsus on seletatav sellega, et tehnoloogilised seadmed, veebilehed ja mobiilirakendused on need, mis tekitavad suure osa uutest andmetest. Laialdased võimalused andmeteadlastele on tingitud ka sektori suurest kasvust. Turundusteenused vajavad andmeteadlasi, sest reklaamindus on digitaliseerunud ja turundajatel on ligipääs suurele hulgale andmetele, mida saab kasutada uute sihtgruppide leidmiseks, kasutatavate meetodite tõhususe hindamiseks ja pakumiste personaliseerimiseks. Finantssektorit võib pidada üheks esimestest majandusharudest, kus hakati rakendama andmeanalüüsi ärilistel eesmärkidel, mida näitab ka termini *business intelligence* ehk ärianalüüsi levik sajandi esimesel kümnendil. Finantsasutused kasutavad andmeanalüüsi näiteks pettuste tuvastamiseks, riskijuhtimiseks, turgude analüüsiks ja erinevateks prognoosideks statistiliste mudelite kaudu. Suurim muutus suurandmetega seoses on ka reaalaajaliste andmete kaasamine analüüsi.

Järgnev lõik põhineb artiklil [10]. Edukaid näiteid andmeteaduse rakendustest võib leida ka muudest valdkondadest. Lennukimootorite tootja Rolls-Royce viib läbi põhjalikud analüüsid uute mootorite välja töötamisel: enne uue toote turule toomist jooksutatakse mitmeid simulatsioone, iga neist toodab kümneid terabaite andmeid. Neid andmeid analüüsides tullakse järeldusele, kas toode on sobilik turule viimiseks või mitte. Ka peale müüki tekitavad mootorid andmeid, et pakuda klientidele kiiret tuge. Sensorid mootorite küljes saadavad reaal-ajas andmeid tagasi ettevõttele, seeläbi on kohe näha muutuseid ja anomaaliaid ning töötajad saavad edastada nõuandeid või hoiatusi klientidele. Selline lähenemine veadiagnostikale ja – ennetusele on aidanud kulusid märgatavalt vähendada. Ameerika jaekaubanduskett avastas andmeanalüüsi abiga, et lisaks tavapärasele varustusele nagu taskulambid suureneb veel osade toodete müük enne tormihoiatusi nagu näiteks maasikakeeksid ja õlu. Seetõttu suurendasid nad enne järgmist tormi nende toodete varu poodides ning enne orkaani saabumist ostetigi neid tooteid suuremates kogustes nagu oodatud. Ka Walmart kasutab andmete reaalaajalist analüüsi müügi jälgimiseks, seeläbi suudeti kiirelt tuvastada, et toode oli unustatud riiulitele panna kui märgati, et osades poodides ei müüdud millegipärast mitte ühtegi vastavat toodet.

Ka Eestis on näiteid andmeanalüüsi kasutamisest ettevõtetes. Riiklik elektritootja Eesti Energia kasutab statistilisi mudeleid, et prognoosida elektritarbimist, eesmärgiga vähendada ületootmist ja sellest tulenevalt kulusid. Statistilisi mudeleid kasutatakse kindlasti ka teistes energeetikavaldkondades, kas siis samuti tootmise prognoosimiseks või muudel eesmärkidel. Näiteks energiasektor on hakanud kasutama biostatistika vallast tuttavat elukestusanalüüsi, et prognoosida varustuse eluiga. Kui eelnevalt oli naftapuurkaevu riketest tingitud seisakuaeg umbes kolm päeva, siis ennetava hooldusega vähendati seda aega kõigest ühe päevani [11].

Artiklites [12-13] on toodud veel näiteid andmeanalüüsi kasutamisest erinevates majandusharudes. Üks näide personaliseeritud reklaamidest tuleb Hotelliketi Red Roof Inn turundusosakonnast, mis kasutab ilmainfot, et leida piirkondi, kus võib katkestatud lendude tõttu olla majutuseta inimesi ning pakub nendel aladel otsingumootorites reklaame. Spordimeeskonnad koguvad andmeid mängijate kohta, lootes ennetada vigastusi. Google teeb koostööd riikliku terviseametiga, et otsingutulemuste põhjal tuvastada piirkondi, kus võivad puhkeda haigushood. Los Angelese politsei koostöös UCLA ülikooliga töötas välja tarkvara, mis põhineb maaväriinate prognoosimisalgoritmil ning pakub vahetust alustavatele politseinikele 10-20 piirkonda, kus toimuvad kuriteod kõige tõenäolisemalt. Üks kaubanduskett muudab oma inventari hindu pea reaajas, vastavalt nõudlusele ja laoseisule. Finantsasutus American Express on välja töötanud mudeli, mis prognoosib tehinguajaloo ja veel 115 tunnuse põhjal, kas firma on kliendist ilma jäämas.

Valim, küsimustik ja uuringu läbiviimine

Uuringu eesmärgiks on välja selgitada vajadus statistiku oskustega töötajate järele, saada teada, millised on ootused nende ettevalmistusele, statistikute väljaõppe parendamine ning ettevõtete vajadustele vastavate statistika-alaste täiendkoolituste väljatöötamine. Vastamise parendamiseks loositakse vastajate vahel välja 45-minutiline statistika-alase koolitus.

Uuringusse kaasati 1000 majanduslikult aktiivset ettevõtet, mis valiti Eesti Statistikaameti poolt 2016. aastal avaldatud majandusüksuste nimekirjast, kokku oli nimekirjas 162 277 üksust. Valim saadi lihtsa juhusliku kihtvaliku rakendamisel, kihtide koostamisel kasutati suurusega võrdelist paigutust, kihistavaks tunnuseks oli ettevõtte tegevusala. Uurijate kõrgendatud huvi tõttu avaliku sektori asutuste vastu kaasati erandina uuringusse 30 üksust kihist avalik haldus ja riigikaitse; kohustuslik sotsiaalkindlustus. Lootes saada informatiivsemaid vastuseid ei kaasatud uuringusse üksuseid järgmistelt tegevusaladelt: kunst, meelelahutus ja vaba aeg ja muud teenindavad tegevused. Lisaks jäeti välja ka koolid ja muud haridusasutused ning korteri, aiandus- ja muud ühistud. Viimaseks ei kaasatud uuringusse töötajateta ettevõtteid, jättes välja esialgsest valimist ettevõtted, mis ei olnud tasunud eelmises (2016. aasta IV) kvartalis tööjõumakse Creditinfo Eesti AS andmetel. Välja langenud ettevõtted asendati uute juhuslikult valitud ettevõtetega ning protsessi korrati, kuni saadi lõplik valim. Kontaktmeilid küsimustiku edastamiseks saadi samuti Creditinfo Eesti kaudu, kuid umbes 5% kirjadest ei jõudnud kindlasti soovitud aadressaadini ilmselt aegunud kontaktinfo tõttu, millele viitasid peale küsimustiku välja saatmist laekunud automaatteavituskirjad. Neid ettevõtteid valimis ei asendatud. On võimalik, et kontaktmeilide hulgas võis olla veel hüljatud meiliaadresse.

Kirjad koos küsimustikuga saadeti välja kolmel korral: esimest korda 11. aprillil ning kirjas oli märgitud vastuste kogumise ajaks kaks nädalat. Teine kiri saadeti vahetult enne kahe nädala möödumist meeldetuletuseks. Viimane kiri saadeti nädal peale seda ning kirjas seletati, et vastuste kogumist jätkatakse nädalaks, et saada rohkem infot. Esimese vooru jooksul saadi 69 vastust, millest 33 tulid peale meeldetuletuskirja, teise vooru jookul lisandus 9 vastust ehk kokku oli vastamismääraks 7,9 %.

Küsimustik paluti täita personaliotsuste või tööülesannete jaotuse eest vastutaval isikul.

Ankeedi esimesed küsimused uurivad arvamust statistiku ametikohast: kas vastaja peab ametit prestiižseks, vajalikuks ja kui suur võiks olla statistiku palk.

Asutuse analüüsitegevuste kohta esitatakse järgnevad küsimused:

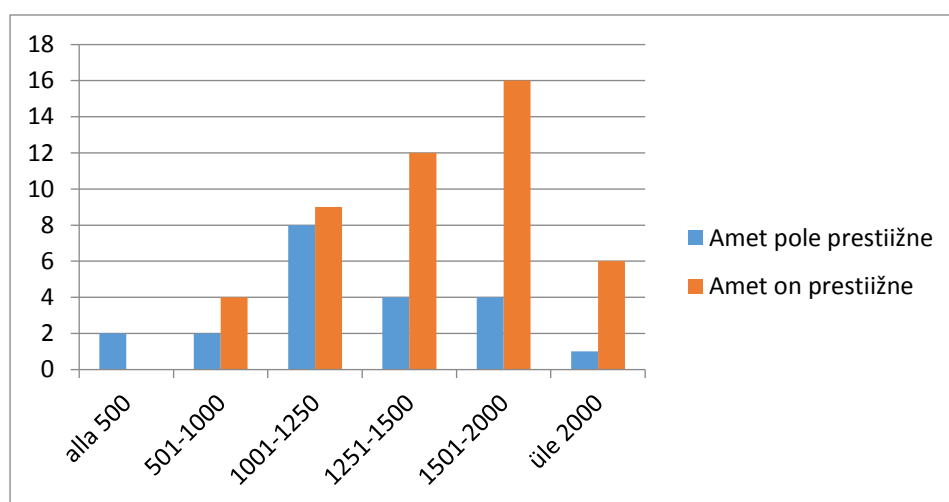
- kas teie asutuses kogutakse igapäevase tegevuse kohta andmeid,
- kas asutuses tehakse statistilist analüüsi,
- kas tulemusi kasutatakse otsuste langetamisel,
- kas analüüsi viib läbi oma asutuses töötav statistik või ostetakse teenus statistilist analüüsi pakkuvalt firmalt,
- mitu statistikut asutuses töötab ning mis väljaõppe on nad saanud,
- milliseid andmeanalüüsi meetodeid kasutatakse asutuses,
- millised oskused on enim hinnatud,
- millist tarkvara kasutatakse statistiliseks analüüsiks ja miks.

Taustinfona on küsitud tegutsemisala, ettevõtte suurust, peakontori asukohta ja kas tegu on era- või avaliku sektori asutusega. Lisaks uuriti, millised oleksid vastajate soovid ja ootused statistika-alase koolituse suhtes.

Kirjeldav analüüs

Statistilise kompetentsi kohta käivad küsimused

Ankeedile laekus 77 vastust. Statistiku eriala hinnangu sektsioonis vastas esimesele küsimusele „Kas selliste oskustega inimene on tööturul vajalik?“ 74 inimest, kusjuures 66 vastas jaatavalt ning 8 eitavalt. Küsimusele „Kas statistiku/analüütiku amet on teie arvates prestiižne?“ vastas jaatavalt 50 ettevõtet, eitavalt 25. Küsimusele „Milline võiks olla teie arvates statistiku/analüütiku brutopalk?“ pakuti kõige enam vahemikku 1501-2000 eurot (20 inimest), teised levinumad arvamused olid 1001-1250 eurot 17 vastusega ning 1251-1500 eurot 16 vastusega. Võrdluseks võib märkida, et 2017. aasta I kvartali keskmine brutopalk oli 1153 eurot [14]. Järgneval joonisel on toodud vastused palga kohta vastavalt sellele, kas vastanud hindasid ametit prestiižseks.



Joonis 1. Hinnang statistiku palgale sõltuvalt küsimuse 2 vastusest.

Järgmise sektsiooni „Üldandmed“ küsimustele vastasid kõik 77 ankeedile vastanut, vastused jagunesid järgnevalt: küsimusele „Kas teie asutuses kogutakse igapäevase tegevuse kohta andmeid?“ vastas jaatavalt 34 ja eitavalt 43 inimest, küsimusele „Kas teie asutuses tehakse statistilist analüüsi?“ vastas jaatavalt 36 ja eitavalt 41 inimest ning küsimusele „Kas teie asutuses kasutatakse statistilise analüüsi tulemusi tegevuste parendamiseks ja/või otsuste langetamiseks?“ vastas jaatavalt samuti 36 ja eitavalt 41 inimest. Tabelis 1 on lisaks nendele arvudele toodud ka tulemused, mille hulgast on välja jäetud ühe töötajaga ettevõtted (tulbad tähistatud tärniga), eelnevate küsimuste vastuseid seda tüüpi ettevõtted ei kallutanud. Tabelist

2 on näha nende kolme küsimuse vastuste ühisjaotus. Selgus, et on väike hulk ettevõtteid, kus viiakse läbi statistilist analüüsi ning kasutatakse selle tulemusi otsuste langetamisel ilma, et need ettevõtted andmeid koguks. Tegevusalalt olid kõik need ettevõtted erinevad, kuid suuruse poolest sarnased, töötajate hulk jäi 1-15 piiresse.

Tabel 1. Andmete kogumise, analüüsi ja selle rakendamise kohta käivate küsimuste vastused.

Küsimus	Jah	Ei	Jah*	Ei*
Kas teie asutuses kogutakse igapäevase tegevuse kohta andmeid?	43	34	41	20
Kas teie asutuses tehakse statistilist analüüsi?	41	36	39	22
Kas teie asutuses kasutatakse statistilise analüüsi tulemusi tegevuste parendamiseks ja/või otsuste langetamiseks?	41	36	38	23

Tabel 2. Andmete kogumise, analüüsi ja selle rakendamise kohta käivate küsimuste vastuste ühisjaotus.

Kas teie asutuses kogutakse igapäevase tegevuse kohta andmeid?	Ei				Jah			
Kas teie asutuses tehakse statistilist analüüsi?	Ei		Jah		Ei		Jah	
Kas teie asutuses kasutatakse statistilise analüüsi tulemusi tegevuste parendamiseks või otsuste langetamiseks?	Ei	Jah	Ei	Jah	Ei	Jah	Ei	Jah
	40.7%	1.16%	2.33%	8.14%	3.59%	2.33%	3.49%	38.37%

Järgmistel on loendatud ettevõtete tegutsemisalad, kus rohkem kui pooltes vastanud ettevõtetes viidi läbi statistilist analüüsi:

- Avalik haldus ja riigikaitse, kohustuslik sotsiaalkindlustus (analüüsi teostati kõikides vastanud asutustes),
- Elektrienergia, gaasi, auru ja konditsioneeritud õhuga varustamine
- Hulgi- ja jaekaubandus, mootorsõidukite ja mootorrattate remont
- Info ja side
- Tervisehoid- ja sotsiaalhoolekanne

Üllatav oli finants- ja kindlustustegevuse puudumine sellest loetelust, tegu võib olla väikese valimi juhuslikkusega (tegevusalalt 6 ettevõtet valimis).

Statistilist analüüsi viib asutustes läbi enamasti astutuses töötav statistik (25 vastust), 4 ettevõtet vastanutest ostavad teenust sisse ning 3 vastasid, et kasutavad mõlemat varianti.

Statistikute arvu ettevõttes uuriva küsimuse vastused on toodud alljärgnevas tabelis.

Tabel 3. Statistikute arv ettevõttes.

Statistikute arv	Vastanuid
1	12
2	3
3	2
4	1
6	1
10	1
> 10	1

Statistikute väljaõpet uurivast küsimusest selgub, et väljaõpped ning haridused on erinevad, kuid kõige levinum on akadeemiline kraad majandusteaduste vallast. Teistest rohkem leidis ka sotsiaalteaduste või IT-kraadiga statistikuid. Vaid kahes ettevõttes olid statistikud matemaatilise statistika kraadiga. Kõik vastused küsimusele on välja toodud Lisade sektsioonis. Küsimuse „Milliseid järgnevas loetelus toodud analüüse teie asutuses tehakse?“ vastused on toodud tabelis 3, küsimusele vastas 32 inimest. Tabelis on lisaks välja toodud, millise valdkonna ettevõtted vastavat analüüsimeetodit sagedamini kasutavad.

Tabel 4. Asutustes kasutatavad analüüsid

Analüüsi ülesanded	Vastanuid	Valdkonnad, kus rakendatakse
Analüüsi tulemuste tõlgendamine ja esitamine	20	Avalik haldus ja riigikaitse, kohustuslik sotsiaalkindlustus; hulgi- ja jaekaubandus, mootorsõidukite ja mootorrataste remont
Andmete kogumine, haldamine, vigade otsimine ja kodeerimine	20	Avalik haldus ja riigikaitse, kohustuslik sotsiaalkindlustus; hulgi- ja jaekaubandus, mootorsõidukite ja mootorrataste remont
Kirjeldav statistika	19	Avalik haldus ja riigikaitse, kohustuslik sotsiaalkindlustus;
Andmete visualiseerimine joonistega	17	Avalik haldus ja riigikaitse, kohustuslik sotsiaalkindlustus;
Uuringu planeerimine, küsimustiku koostamine, valimimahu hindamine	11	Avalik haldus ja riigikaitse, kohustuslik sotsiaalkindlustus; Muud teenindavad tegevused
Andmete mudeldamine	9	Info ja side, Avalik haldus ja riigikaitse, kohustuslik sotsiaalkindlustus;
Andmete imputeerimine	7	Avalik haldus ja riigikaitse, info ja side, muud teenindavad tegevused
Suurandmete analüüs	4	Info ja side

Küsimuse „Millised oskusi hindate statistiku/analüütiku puhul kõige rohkem?“ vastused on toodud allpool tabelis 4, küsimusele vastas 30 küsitletut.

Tabel 5. Enimhinnatud oskused statistikutel.

Statistiku jaoks olulised oskused	Vastanuid
Statistiliste tulemuste esitamine arusaadavalt mittestatistikutele	20
Suurte andmehulkade analüüsimine	19
Teie valdkonnas kasutatavate statistiliste meetodite, sh mudelite valdamine	16
Visualiseerimisoskused	15
Statistiliste tulemuste interpreteerimine	14
Suurte andmehulkade haldamine	13
Teie valdkonnaspetsiifilised teadmised, ilma milleta andmeanalüüsi tegemine on võimatu	11

Kasutatava tarkvara kohta uurivast küsimusest selgus, et kõige levinumaks programmiks on Microsoft Excel: 29 29-st vastanust vastas, et ettevõtte kasutab Excelit andmeanalüüsiks, ainult Excelit kasutas neist 19 ettevõtet. Kõik vastused on toodud järgnevas tabelis 5.

Tabel 6. Asutustes kasutatavad analüüsitarkvarad

Tarkvara	Kasutajaid
Excel	29
R	4
SPSS	4
Python	4
Directo	1
Qlikview	1
STATISTICA	1
Tableau	1
Matlab	1

Tarkvara valikut põhjendati kõige enam mugavusega (15 vastust). Nii hinna kui ka tuntusega põhjendas tarkvara valikut 7 ettevõtet, statistiku soovi põhjal olid valiku teinud ainult 4 ettevõtet. Nendest neljast kasutati kahes ettevõttes Excelit, teises kahes R-i ja Pythonit.

Ettevõtetest, kus statistikut palgal ei olnud, vastasid eitavalt küsimusele „Kas olete oma statistiku/analüütiku palkamisele mõelnud?“ 43 ning jaatavalt 2 vastanut. Eitavalt vastanutest põhjendasid 19 neist oma valikut variandiga „Statistilisel analüüsil ei tundu olevat asutuse eesmärkide seisukohalt lisaväärtust“ ning 7 vastasid, et tellivad analüüsi mõnelt väliselt ettevõttelt. Vabavastustest toodi välja enamasti ettevõtte väike suurus. Tabelis 7 on toodud samad tulemused sõltuvalt ettevõtte suurusest.

Tabel 7. Küsimuse "Miks te ei ole mõelnud statistiku/analüütiku palkamise peale?" vastused

	Ühe töötajaga ettevõtted	2-15 töötajaga ettevõtted	rohkem kui 15 töötajaga ettevõtted
Statistilisel analüüsil ei tundu olevat asutuse eesmärkide seisukohalt lisaväärtust	6	11	2
Tellime vajadusel statistilise analüüsi antud teenust pakkuvalt ettevõttelt	2	3	2

Küsimusele „Miks te statistikut/analüütikut pole siiani palganud“ saadi 2 vastust, variandid „Töömaht on statistiku/analüütiku palkamiseks liiga väike“ ning „Tellime vajadusel statistilise analüüsi antud teenust pakkuvalt ettevõttelt“ said kumbki ühe vastuse. Sobiva kandidaadi puudumist ei pidanud kumbki vastanu probleemiks. Küsimus, mis proovis uurida seniste kandidaatide nõrkusi kahjuks vastuseid ei saanud.

Vastanute taustinfo

Järgnevalt on toodud vastused küsimustele, mis pärisid infot vastanud ettevõtte kohta. Tuleb märkida, et vastanute ettevõtte liigitus ei pruugi kattuda äriregistri Eesti Majanduse Tegevusalade Klassifikaatori, mille järgi valiti valimi arv, liigitusega.

Tabel 8. Vastanute arv ja vastustemäärad ettevõtte tegevusala järgi

Ettevõtte liigitus	Vastanuid	Arv valimis	Vastamismäär
Avalik haldus ja riigikaitse; kohustuslik sotsiaalkindlustus	8	30	27%
Ehitus	6	110	5%
Elektrienergia, gaasi, auru ja konditsioneeritud õhuga varustamine	1	5	20%
Finants- ja kindlustustegevus	6	40	15%
Hulgi- ja jaekaubandus; mootorsõidukite ja mootorrataste remont	7	180	4%
Info ja side	3	65	5%
Kinnisvaraalne tegevus	3	85	4%
Kunst, meelelahutus ja vaba aeg	5	0	-
Kutse-, teadus- ja tehnikaalne tegevus	6	145	4%
Majutus ja toitlustus	1	30	3%
Muud teenindavad tegevused	11	0	-
Põllumajandus, metsamajandus ja kalapüük	4	50	8%
Tervishoid ja sotsiaaltoetused	6	25	24%
Töötlev tööstus	6	75	8%
Veevarustus; kanalisatsioon, jäätme- ja saastekäitlus	1	5	20%
Veondus ja laondus	3	50	6%

Hoolimata proovist eraldada valimist ühe töötajaga ettevõtteid, sattus neid siiski vastanute hulka. Võib arvata, et tõenäoliselt oli ka selliste ettevõtete vastamismäär suurem: kuna küsitlus saadeti ettevõtte üldkontaktmeilile, siis väikses ettevõttes jääb kirja kohale jõudmise ja ankeedile vastamise vahele vähem vaheetappe, milletõttu küsimustik lõpuks täitmata võib jääda (kuigi me ei tea tegelikku valimisse sattunud väikeettevõtete arvu). Enamus sellistest ettevõtetest tegutsesid järgmistes valdkondades: finants- ja kindlustustegevus; hulgi- ja jaekaubandus, mootorsõidukite ja mootorrataste remont; kutse-, teadus- ja tehnikalane tegevus; muud teenindavad tegevused.

Tabel 9. Vastanud ettevõtete suurus

Töötajate arv	Vastanuid
1	16
2-4	12
5-15	25
16-50	7
51-100	3

Uuriti ka ettevõtte peakontori asukohta.

Tabel 10. Ettevõtte peakontori asukoht

Haapsalu/Läänemaa	1
Harjumaa	6
Jõhvis	1
Kuressaare	3
Pärnumaa/Pärnu	2
Tallinn	41
Tartumaa/Tartu	12
Valgamaa	1
Viljandi	2
Võru	1

Küsitlust täitis kõige enam kas ettevõtte juhataja (19 vastust, neist kümnes oli juhataja ettevõtte ainutöötaja) või juhatuse liige (10 vastust). 4 korral täitis ankeeti personalijuht, 10

küsimustikku täitis mõni muu keskastmejuht, lisaks veel muud erinevad vastused (kokku 57 vastust küsimusele, vt. Lisa 1).

Koolituse kohta kogutud info

Viimane osa küsitlusest uuris vastanute ootusi statistika-alase koolituse suhtes. Üks kolmandik vastanutest ehk 25 inimest ütlesid, et oleksid huvitatud oma töötajate saatmisest andmeanalüüsi koolitusele. Nende hulgas oli neli ettevõtet, kus ei tegeletud analüüsiga, selles grupis oli suurem huvi uuringu planeerimise vastu ja kirjeldava statistika vastu. Teemasid, mida kõik vastanud sooviksid koolitusel näha, on toodud järgnevas tabelis.

Tabel 11. Koolituse teemad.

Koolitusel käsitletav teema	Vastanuid
Kirjeldav statistika (tabelid, arvkarakteristikud)	19
Andmete visualiseerimine joonistega	18
Analüüsi tulemuste tõlgendamine ja esitamine	18
Andmete mudeldamine	15
Uuringu planeerimine, küsimustiku koostamine, valimimahu hindamine	12
Andmete imputeerimine (puuduvate väärtuste ennustamine)	11
Andmetest vigade otsimine	10
Suurandmete analüüs (masinõpe, peakomponentanalüüs, faktoranalüüs, tugivektorid jne)	7
Andmete sisselugemine ja manipuleerimine	6

21 22-st küsimusele “Mis võiks olla koolitusel kasutatav tarkvara?” vastanust soovisid, et koolitus viidaks läbi Exceliga, muudest tarkvaradest oli rohkem pakutud R-i ja SPSS-I (4 ja 3 vastust). Excel oli ainuke valik nende ettevõtete seas, kus statistilist analüüsi läbi ei viida (tulenevalt vastusest küsimusele 5). 6 vastanut eelistasid, et koolituse vormiks võiks olla auditoorne loeng koos praktiliste ülesannete lahendamisega, 4 soovisid ainult praktiliste näiteülesannete lahendamist (soovi korral asutuse enda andmetega). Kolmandat varianti, mõlemad vormid integreeritud, eelistati kõige rohkem (14 vastanut). Koolituse pikkuseks määrati kõige enam päev (9 vastanut) või paar päeva (8 vastanut). Koolituse maksumuse kohta ei valitud ühtegi vastusevarianti rohkem kui kord või ei osatud küsimusele vastata.

Kokkuvõte

Käesoleva bakalaureusetöö eesmärgiks oli uurida vajadust statistiku või analüütiku oskuste järgi Eestis tegutsevates ettevõtetes. Esimeses peatükis tutvustati analüütikal põhinevate äriotsuste ajalugu ning nende kasutamise vajaduse suurenemist ülemaailmselt. Toodi ka mõned näited sarnastest küsitlustest mujal maailmas. Teises peatükis kirjeldati valimi moodustamist, ankeedi küsimusi ja uuringu läbiviimist. Kolmas peatükk koosneb uuringu tulemuste kirjeldavast statistikast.

Kuigi küsitluse vastamismäär jäi oodatust madalamaks, on siiski võimalik mõningad järeldused teha. Arvamus statistiku ametikohast on hea ning üle poolte uuritud asutustest, kus on palgal rohkem kui üks töötaja, kogub andmeid ja kasutab neid analüüsiks ja otsuste või tegevuste parendamiseks. Levinumad analüüsivormid Eesti ettevõtetes olid analüüsi tulemuste tõlgendamine ja esitamine, andmete kogumine, haldamine, vigade otsimine ja kodeerimine, kirjeldav statistika ning andmete visualiseerimine joonistega. Kõige hinnatumaks oskuseks kujunes analüüsi tulemuste esitamine mittestatistikutele. Analüüsitarkvaradest on kõige laialdasemalt kasutuses Excel, välja võib tuua lisaks ka R-i, Pythoni ja SPSS-i. Kolmandik vastanutest oleksid huvitatud oma töötajate saatmisest statistika-alasele täienduskoolitusele. Soov oli näha koolitusel järgnevaid teemasid: kirjeldav statistika, andmete visualiseerimine joonistega, analüüsi tulemuste tõlgendamine ja esitamine, andmete mudeldamine. Koolituse vormiks valiti pikkuseks valiti kõige enam päev või paar päeva ning koolituse toimumise enim valitud variandiks oli kaks teist pakutud vormi (ainult praktiliste näiteülesannete lahendamine, soovi korral asutuse enda andmetega ja auditoorne loeng koos praktiliste ülesannete lahendamisega).

Probleemseks osutus valimi moodustamisel ühe töötajaga ettevõtete välja jätmine valimist, mis oli määratud eelmise kvartali tasutud tööjõumaksude tasumise järgi. Üheks madala vastamismäära põhjustest võivad olla aegunud avalikud andmed, millele viitasid ligikaudu 50 automaatvastuskirja teavitusega mitteeksisteerivast meiliaadressist. Edaspidiste uuringu raames võiks kaaluda valimi tellimist.

Kasutatud kirjandus

- [1] Gordon, J., Perrey, J., Spillecke, D(2013). *Big Data, Analytics And The Future Of Marketing And Sales*. <https://www.forbes.com/sites/mckinsey/2013/07/22/big-data-analytics-and-the-future-of-marketing-sales/#529717185587> (30.04.2017).
- [2] More Students Earning Statistics Degrees, But Not Enough To Meet Surging Demand for Statisticians. *This Is Statistics* (2015).
<http://thisisstatistics.org/more-students-earning-statistics-degrees-but-not-enough-to-meet-surging-demand-for-statisticians/> (30.04.2017).
- [3] US Census Bureau. PINC-01. Selected Characteristics of People 15 Years and Over, by Total Money Income, Work Experience, Race, Hispanic Origin, and Sex.
<https://www.census.gov/data/tables/time-series/demo/income-poverty/cps-pinc/pinc-01.html> (30.04.2017).
- [4] Fisher, C (2016). LinkedIn Unveils The Top Skills That Can Get You Hired In 2017, Offers Free Courses for a Week. *LinkedIn Official Blog*.
<https://blog.linkedin.com/2016/10/20/top-skills-2016-week-of-learning-linkedin> (30.04.2017).
- [5] Mulvey, T., Esen, E., Schramm, J (2016). *Jobs of the Future: Data Analysis Skills*.
<https://www.shrm.org/hr-today/trends-and-forecasting/research-and-surveys/Documents/Data-Analysis-Skills.pdf> 30.04.2016 (30.04.2017).
- [6] Morris, R., Trukowski, B.J. (2003). The evolution of storage systems. *IBM System Journals*, 42, 205-217.
- [7] Nate Silver: What I need from statisticians. *Statistic Views* (2013).
<http://www.statisticviews.com/details/feature/5133141/Nate-Silver-What-I-need-from-statisticians.html> (30.04.2017).
- [8] Press, G. (2013). Data Science: What's The Half-Life Of A Buzzword? *Forbes*
<https://www.forbes.com/sites/gilpress/2013/08/19/data-science-whats-the-half-life-of-a-buzzword/#443451dd7bfd> (30.04.2017).
- [9] Burtch, L. (2016). Salaries of Data Scientists. *Burtch Works Study*.
http://www.burtchworks.com/files/2016/04/Burtch-Works-Study_DS-2016-final.pdf (30.04.2017).
- [10] Marr, B. (2016). The Most Practical Big Data Use Cases Of 2016. *Forbes*.

<https://www.forbes.com/sites/bernardmarr/2016/08/25/the-most-practical-big-data-use-cases-of-2016/> (30.04.2017).

[11] Harpham, B. (2016). How data science is changing the energy industry. *CIO Online*.
<http://www.cio.com/article/3052934/big-data/how-data-science-is-changing-the-energy-industry.html> (30.04.2017).

[12] Laskowski, N. (2013) . Ten big data case studies in a nutshell. *SearchCIO*.
<http://searchcio.techtarget.com/opinion/Ten-big-data-case-studies-in-a-nutshell> (30.04.2017).

[13] Petersen, R. (2016). 37 Big Data Case Studies with Big Results. *{grow}*.
<https://www.businessesgrow.com/2016/12/06/big-data-case-studies/> (30.04.2017).

[14] Statistikaamet. Keskmine brutokuupalg. <https://www.stat.ee/stat-keskmine-brutokuupalk> (07.01.2017).

Lisad

Lisa 1. Vabavormis vastused

Küsimusele “Mitu statistikut/analüütikut teie ettevõttes töötab” laekusid järgnevad vabas vormis vastused:

- Need inimesed on muu ala professionaalid, näiteks tootejuht
- Liiga suur asutus, et neid lugeda
- Ühtegi, ise teeme
- Erinevad isikud tegelevad analüüsiga, vastavalt tööülesannetele

Küsimuse “Mis haridusega on teie statistik/analüütik(ud)?” kõik laekunud vastused:

- matemaatilise statistika PhD, MSc, epidemioloogia MSc
- ei tea :)
- sotsioloog
- statistik
- jurist
- mehhaanika, TTÜ haridusega tavaliselt
- majandusteaduste bakalaureus
- Ei oska haridustaset täpsustada, sisuliselt õpime koha peal jooksvalt seda, mida vaja - pole eriti keeruline tase.
- majandusteaduste bakalareus
- erinevad kõrgharidused, vastavalt isikule
- magister, majandus, aga ka muu valdkonna ekspert
- erinevad
- keemia
- majandusteaduste bakalaureus
- bakalaureus, magister, erinevad erialad
- Statistika, infotehnoloogia
- sotsiaalteaduse magister, majandusteaduse magister, sotsiaalteaduse bakalaureus, õigusteaduse bakalaureus, haldjusjuhtimise bakalaureus, matemaatilise statistika bakalaureus
- Erinevad kuid ükski pole statistik
- bakalaureus
- Sotsiaalteaduste bakalaureus

- iseenese tarkuse
- IT magister
- pole erialast haridust
- riigiteaduste bakalaureus, ärikorranduse rakenduskõrgharidus, ühiskonnaõpetuse õpetaja, personalijuhtimise bakalaureus
- maj magister
- Finantsharidusega töötaja rakenduskõrg

Küsimusele “Miks te olete valinud just selle [analüüsi] tarkvara?” laekunud vabavastused:

- Oli kättesaadav
- Kättesaadav, põhjalikumad andmeanalüüsi pole vaja
- Puudub muu oskusteave
- Tarkvara võimekus analüüsi toetada, mugavus ja hind

Küsimusele “Miks te ei ole mõelnud statistiku/analüütiku palkamise peale?” laekunud vabavormis vastused:

- ei ole vaja, niipalju kui vaja suudan ise
- Ettevõtte liiga väike
- Juhatusel on teadmised statistilisest analüüsist, vajadusel koostab statistilisi analüüse.
- meie tegevusalal ei ole selle mõtet
- mikroettevõtte, ei ole vajadust ega võimalust
- Pole olnud vajadust
- pole oluline
- pole vajadust olnud
- Põhitegevus baseerub teistel alustel, analüüsitavad andmemahud ei ole suured.
- saan ise hakkama. bakalaureuse kraad oli sotsioloogias ja statistika alused selged
- Statistilisel analüüsil ei tundu olevat asutuse eesmärkide seisukohalt lisaväärtust, Puudub vajadus olemasolevaid või ka tulevasi andmeid statistiliselt analüüsida.
- Usun , et oman ülevaadet ilma spetsialistita
- väike ettevõtte
- Väikefirma ja finantsiliselt pole võimalik lubada
- Ühemehefirma

Küsimusele “Kes täitis ankeedi” kõik laekunud vastused:

- analüütik
- Andmeanalüüsi teenusepakkuja
- boss
- CTO
- hambaarst
- juh.liige
- Juhataja
- Juhatuse esimees
- juhatuse liige
- juhatuse liige, osanik
- keraamik
- keskastme juht
- Koolitaja
- koristaja-peadirektor
- koristajast juhatuse liikmeni
- Korrapidaja
- Kõnekeskuse juht
- Müügiassistent
- nõunik
- Osakonnajuhataja
- Peraamatupidaja
- perearst
- Personali- ja haldusjuht
- Personalibüroo juhataja
- personalijuht
- personalispetsialist
- raamatupidaja
- Spetsialist / juhatuse liige
- tegevjuht
- tislér
- tootmisjuht

- Uuringu projektijuht
- võtmekliendihaldur
- Ärianalüüsi osakonna juhataja

Lisa 2. Küsimustik



TARTU ÜLIKOOL
matemaatika ja statistika instituut

Igas ettevõttes on järjest enam andmeid nii tegutsemisprotsessi kui klientide kohta. Selle informatsiooni kasutamisel ettevõtte tegevuse parendamiseks on abiks analüütikud.

Koostöös Tartu Ülikooli Matemaatika ja statistika instituudiga on käesoleva uuringu eesmärgiks hinnata statistilise/analüütilise kompetentsi vajadust Eestis tegutsevates ettevõtetes. Tahame välja selgitada praeguse olukorra ning vajaduse statistikute/analüütikute järele. Millised on ootused nende ettevalmistusele ja oskustele. Loodame uuringust saada suuniseid statistikute koolituse arendamiseks nii, et see vastaks paremini ettevõtete vajadustele.

Küsitlus on anonüümne. Küsitluse täitmine võtab aega 5-10 minutit. Küsimuste korral pöörduge aadressil kaldermaarand@gmail.com.

Kõigi vastajate vahel, kes soovivad, loositakse välja 45 minutiline statistika-alane konsultatsioon.

Statistiku eriala hinnang

Matemaatiline statistika tähendab tööd informatsiooniga: selle kogumist, töötlemist, analüüsimist ja lahtimõtestamist. Matemaatilise statistika eriala lõpetanu oskab kasutada arvutit suure ning keeruka struktuuriga andmete analüüsimisel. Tänapäeva statistikud on intensiivsed arvutikasutajad, kes vajaduse korral ka ise tarkvara loovad, kuid erinevalt informaatikutest on neil siiski esikohal analüüsi meetodika ja tulemuste tõlgendamine.

1. Kas selliste oskustega inimene on tööturul vajalik?

- ☐ Jah
☐ Ei

2. Kas statistiku/analüütiku amet on teie arvates prestiižne?

- ☐ Jah
☐ Ei

3. Milline võiks olla teie arvates statistiku/analüütiku brutopalk?

- ☐ alla 500
- ☐ 500-1000
- ☐ 501-1000
- ☐ 1001-1250
- ☐ 1251-1500
- ☐ 1501-2000
- ☐ üle 2000
- ☐ Other: _____

Üldandmed

4. Kas teie asutuses kogutakse igapäevase tegevuse kohta andmeid? *

Näited andmetest firma tegevuse kohta: 1) reisijateveo teenust pakkuv asutus kogub andmeid reisijate arvu, marsruudi, sooduspiletide ostjate jne kohta; 2) toidupood kogub kliendikaardi abil andmeid ostjate ostueelistuste kohta; 3) ministeerium kogub andmeid elamisloa taotluse esitanud isikute kohta.

- ☐ Jah
- ☐ Ei *Peale seda sektsiooni liikuge küsimus 19 juurde.*

5. Kas teie asutuses tehakse statistilist analüüsi? *

Statistilise analüüsi alla loetakse järgnevad tegevused: andmete teisendamine, haldamine, andmete tabeldamine, andmete visualiseerimine graafikutena, andmete kirjeldamine arvkarakteristikutega (keskmine, standardhälve jne), andmete mudeldamine, tulemuste tõlgendamine jne.

- ☐ Jah
- ☐ Ei *Peale seda sektsiooni liikuge küsimus 14 juurde*

6. Kas teie asutuses kasutatakse statistilise analüüsi tulemusi tegevuste parendamiseks ja/või otsuste langetamiseks? *

- ☐ Jah
- ☐ Ei

Kes statistilist analüüsi läbi viib?

7. Kes statistilist analüüsi läbi viib?

- ☐ Oma asutuses töötav statistik/analüütik
- ☐ Ostate kompetentsi sisse statistilist analüüsi pakkuvast firmast *Liikuge küsimus 14 juurde*
- ☐ Mõlemad eelpool toodud variandid

Statistiku/analüütiku(te) väljaõpe

Siia palun kirjutage kõikide statistikute/analüütikute, kes teie asutuses töötavad, formaalne haridus.

8. Mitu statistikut/analüütikut teie asutuses töötab?

Kui teie asutuses töötab rohkem kui viis statistikut/analüütikut, siis palun kirjutage täpne arv variandi "Muu".

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ Other: _____

9. Millise väljaõppega on teie statistik/analüütik(ud)?

Näiteks majandusteaduste magister, matemaatilise statistika bakalaureus jne.

Kasutatavad meetodid

10. Milliseid järgnevas loetelus toodud analüüse teie asutuses tehakse?

Valige kõik sobivad.

- ☐ Uuringu planeerimine, küsimustiku koostamine, valimimahu hindamine
- ☐ Andmete kogumine, haldamine, vigade otsimine ja kodeerimine
- ☐ Kirjeldav statistika (tabelid, arvkarakteristikud)
- ☐ Andmete visualiseerimine joonistega
- ☐ Andmete imputeerimine (puuduvate väärtuste ennustamine)
- ☐ Suurandmete analüüs (masinõpe, peakomponentanalüüs, faktoranalüüs, tugivektorid jne)
- ☐ Andmetest mustrite/anomaaliate otsimine
- ☐ Andmete mudeldamine
- ☐ Analüüsi tulemuste tõlgendamine ja esitamine
- ☐ Other: _____

11. Millised oskusi hindate statistiku/analüütiku puhul kõige rohkem?

Valige kõik sobivad.

- ☐ Avalikest andmebaasidest andmete kättesaamine
- ☐ Suurte andmehulkade haldamine
- ☐ Suurte andmehulkade analüüsimine
- ☐ Visualiseerimisoskused
- ☐ Teie valdkonnas kasutatavate statistiliste meetodite, sh mudelite valdamine
- ☐ Statistiliste tulemuste interpreteerimine
- ☐ Statistiliste tulemuste esitamine arusaadavalt mittestatistikutele
- ☐ Teie valdkonnaspetsiifilised teadmised, ilma milleta andmeanalüüsi tegemine on võimatu
- ☐ Other: _____

Tarkvara

12. Millist tarkvara teie asutus kasutab analüüsi läbiviimiseks?

Valige kõik sobivad.

- ☐ EXCEL
- ☐ R
- ☐ SPSS
- ☐ STATA
- ☐ SAS
- ☐ Mplus
- ☐ Matlab
- ☐ STATISTICA
- ☐ Python
- ☐ Other: _____

13. Miks olete valinud just selle tarkvara?

Valige kõik sobivad.

- ☐ Mugavus
- ☐ Tuntus
- ☐ Hind
- ☐ Statistiku/analüütiku soov
- ☐ Other: _____

*Liikuge küsimus 19
juurde*

Andmeanalüütik asutuses

14. Kas olete oma statistiku/analüütiku palkamisele mõelnud?

- ☐ Jah
- ☐ Ei

*Liikuge küsimus 16
juurde*

Palkamine

15. Miks te ei ole mõelnud statistiku/analüütiku palkamise peale?

Valige kõik sobivad.

- ☐ Statistilisel analüüsil ei tundu olevat asutuse eesmärkide seisukohalt lisaväärtust
- ☐ Tellime vajadusel statistilise analüüsi antud teenust pakkuvalt ettevõttelt
- ☐ Other: _____

Liikuge küsimus 19 juurde

Põhjused

16. Miks te statistikut/analüütikut ei ole siiani palganud?

Valige kõik sobivad.

- ☐ Statistikut/analüütikut ise ei ole huvi üles näidanud meie vastu
- ☐ Statistiku/analüütiku palkamiseks ei ole piisavalt rahalisi vahendeid
- ☐ Töömaht on statistiku/analüütiku palkamiseks liiga väike
- ☐ Olemasolevad töötajad on suutnud ise vajalikul määral andmeanalüüsi sooritada
- ☐ Tellime vajadusel statistilise analüüsi antud teenust pakkuvalt ettevõttelt
- ☐ Other: _____

17. Kas sobiva kandidaadi puudumine on üks põhjustest, miks te siinemaani statistikut/analüütikut pole palganud?

- ☐ Jah
- ☐ Ei *Liikuge küsimus 19 juurde*

Oskused

18. Millised oskused on senistel kandidaatidel olnud nõrgad/puudunud?

Valige kõik sobivad.

- ☐ Avalikest andmebaasidest andmete kättesaamine
- ☐ Suurte andmehulkade haldamine
- ☐ Suurte andmehulkade analüüsimine
- ☐ Visualiseerimisoskused
- ☐ Teie valdkonnas kasutatavad statistilised meetodid, sh mudelid
- ☐ Statistiliste tulemuste interpreteerimine
- ☐ Statistiliste tulemuste esitamine arusaadavalt mittestatistikutele
- ☐ Teie valkonnaspetsiifilised teadmised, ilma milleta andmeanalüüsi tegemine on võimatu
- ☐ Other: _____

Taustinfo

19. Millisel alal teie asutus tegutseb? *

- ☐ Põllumajandus, metsamajandus ja kalapüük
- ☐ Mäetööstus
- ☐ Töötlev tööstus
- ☐ Elektrienergia, gaasi, auru ja konditsioneeritud õhuga varustamine
- ☐ Veevarustus; kanalisatsioon, jäätme- ja saastekäitlus
- ☐ Ehitus
- ☐ Hulgi- ja jaekaubandus; mootorsõidukite ja mootorrataste remont
- ☐ Veondus ja laondus
- ☐ Majutus ja toitlustus
- ☐ Info ja side
- ☐ Finants- ja kindlustustegevus
- ☐ Kinnisvaraala tegevus
- ☐ Kutse-, teadus- ja tehnikala tegevus
- ☐ Haldus- ja abitegevused
- ☐ Avalik haldus ja riigikaitse; kohustuslik sotsiaalkindlustus
- ☐ Haridus
- ☐ Tervishoid ja sotsiaaltoolekanne
- ☐ Kunst, meelelahutus ja vaba aeg
- ☐ Muud teenindavad tegevused
- ☐ Kodumajapidamiste kui tööandjate tegevus; kodumajapidamiste oma tarbeks mõeldud eristamata kaupade tootmine ja teenuste osutamine
- ☐ Eksterritoriaalsete organisatsioonide ja üksuste tegevus

20. Kui palju töötajaid teie asutuses on? *

- ☐ 1
- ☐ 2-4
- ☐ 5-15
- ☐ 16-50
- ☐ 51-100
- ☐ Üle 100

21. Kus tegutseb teie asutuse peakontor?

- ☐ Tallinnas
- ☐ Tartus
- ☐ Narvas
- ☐ Pärnus
- ☐ Kohtla-Järvel
- ☐ Other: _____

22. Kas tegemist on erasektori või riigiasutusega?

- ☐ Erasektor
- ☐ Riigiasutus

23. Vastaja enda ametikoht

24. Kas oleksite huvitatud oma töötajate saatmisest andmeanalüüsi koolitusele?

Nii asutuse spetsiifiline üksiku(te) isiku(te) koolitus kui ka laiem statistiliste oskuste arendamise koolitus

- ☐ Jah
- ☐ Ei *Liikuge küsimus 30 juurde.*

Koolituse kohta käiv info

25. Millised teemad võiksid olla käsitletud?

Valige kõik sobivad.

- ☐ Uuringu planeerimine, küsimustiku koostamine, valimimahu hindamine
- ☐ Andmete sisselugemine ja manipuleerimine
- ☐ Andmetest vigade otsimine
- ☐ Kirjeldav statistika (tabelid, arvkarakteristikud)
- ☐ Andmete visualiseerimine joonistega
- ☐ Andmete imputeerimine (puuduvate väärtuste ennustamine)
- ☐ Suurandmete analüüs (masinõpe, peakomponentanalüüs, faktoranalüüs, tugivektorid jne)
- ☐ Andmete mudeldamine
- ☐ Analüüsi tulemuste tõlgendamine ja esitamine
- ☐ Other: _____

26. Koolitusel kasutatav tarkvara võiks olla

Valige kõik sobivad.

- ☐ EXCEL
- ☐ R
- ☐ SPSS
- ☐ STATA
- ☐ SAS
- ☐ Python
- ☐ Other: _____

27. Koolituse toimumise vorm

Sõltumata toimumise vormist oleks koolitusel/(video)loengutes kohal ka juhendajad

- ☐ Auditoorne loeng koos praktilise ülesannete lahendamisega
- ☐ Internetipõhine loeng koos iseseisvate praktiliste ülesannetega
- ☐ Mõlemad vormid integreeritult
- ☐ Ainult praktiliste näiteülesannete lahendamine, soovi korral asutuse enda andmetega
- ☐ Other: _____

28. Kui pikk võiks koolitus olla?

- ☐ Paar tundi kuni pool päeva
- ☐ 1 päev
- ☐ Paar päeva
- ☐ Nädal
- ☐ Paarikuune, iseseisvate koduste ülesannetega
- ☐ Other: _____

29. Kui palju võiks koolitus maksta?

Suur tänu vastamast!

- 30. Kui soovite osaleda konsultatsiooni loosimises, lisage siia kontakt e-mail või saatke kiri e-mailile kaldermaarand@gmail.com.**

- 31. Kui soovite veel midagi lisada, palun kirjutage see siia.**

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, Kalder Maarand,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose „Statistilise kompetentsi vajadus Eestis tegutsevates ettevõtetes“, mille juhendajateks on lektor Mare Vähi, Silva Kasela ja Kristi Läll,

1.1. reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;

1.2. üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.

2. olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.

3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus, 13.02.2018